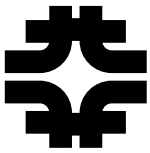


Fermi National Accelerator Laboratory

FERMILAB-Pub-96/053-A

Higher Order Statistics from the EDSGC Survey I: Counts in Cells

Istvan Szapudi et al.

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

January 1996

Submitted to *Astrophysical Journal*

Operated by Universities Research Association Inc. under Contract No. DE-AC02-76CHO3000 with the United States Department of Energy

Disclaimer

FERMILAB-Pub-96/053-A

submitted to ApJ.

Higher Order Statistics from the EDSGC Survey I: Counts in Cells

István Szapudi[1]
Fermi National Accelerator Laboratory
Theoretical Astrophysics Group
Batavia, IL 60510

Avery Meiksin[2] and Robert C. Nichol[3]
University of Chicago
Department of Astronomy & Astrophysics
5640 South Ellis Avenue
Chicago, IL 60637

ABSTRACT

Counts in cells are used to analyse the higher order properties of the statistics of the EDSGC survey. The probability distribution is obtained from an equal area projection point source catalog with massive oversampling in the scale range of $0.015^\circ - 2^\circ$. The factorial moments of the resulting distribution and the s_N's characterizing the non-Gaussian nature of the distribution are extracted. These results are directly comparable to previous results from the APM survey, and to theoretical results from perturbation theory. The deprojected 3D values corresponding to the s_N's are also determined. We find that the 3D values match the scaling relation for strongly nonlinear clustering found in N-body simulations remarkably well.

Subject headings: large scale structure of the universe — methods: numerical

---

[1] E-mail: szapudi@astro1.fnal.gov

[2] Edwin Hubble Research Scientist; E-mail: meiksin@oddjob.uchicago.edu

[3] E-mail: nichol@huron.uchicago.edu

1. Introduction

A leading hypothesis for the origin of the large-scale structure of the distribution of galaxies is that it is a consequence of gravitational instability in an initially homogeneous medium. The $N-$point correlation functions provide a set of statistics suited for quantifying the expected departure from homogeneity of the galaxy distribution under this hypothesis (Peebles 1980). The statistical analyses of recent galaxy catalogs has tended to provide support for this scenario. While the 2–point correlation function has clearly demonstrated the non-Poisson character of the galaxy distribution, it is not a unique test of gravitational instability since it is reproduced by a variety of models for structure formation (Peebles 1993). If gravitational instability dominates the growth of structure, however, then it is possible to predict a relation between the higher-order correlation functions and the 2–point function. In particular, if the structure is hierarchical in nature, as expected in the strongly nonlinear limit, then the N–point functions are symmetrized products of $N-1$ 2–point functions (Peebles 1980), with well-defined theoretical predictions for their amplitudes (Hamilton 1988). In the limit of weakly nonlinear clustering, analytic forms for the amplitudes in analogous relations between spatial averages of the correlation functions have been derived (Juszkiewicz, Bouchet, & Colombi 1993; Colombi *et al.* 1994; Bernardeau 1994a,b).

Angular catalogs offer two advantages over their redshift analogs for measuring higher-order correlations: their large size and their insensitivity to redshift distortions. A disadvantage is that, because they are projections of the galaxy distribution, simplifying assumptions must be made concerning the clustering of galaxies to make the extraction of the higher order correlations practical. Thus the analyses of both types of catalogs are complementary. Measurements of the higher order correlation functions in angular catalogs have supported the form predicted by the hierarchical model. The amplitudes, however, have shown some variance, depending on the method of analysis and the catalog. Szapudi, Szalay, & Boschán (1992) confirmed and refined the estimate of Groth & Peebles (1977) for the 3–point function of the Lick counts (Shane & Wirtanen 1967), although their estimate of the amplitude of the 4–point function falls somewhat below that of Fry & Peebles (1978). Szapudi et al. provide estimates for higher order functions as well. Analyses of the IRAS catalogs have provided even stronger support for the hierarchical model, although the correlations of these infrared-selected galaxies tend to be somewhat weaker than those of their optical counterparts, perhaps reflecting a genuine morphology-dependence in the nature of clustering (Meiksin, Szapudi, & Szalay 1992, Bouchet *et al.* 1993). Recently, the analysis of higher order functions has been extended to the APM catalog (Maddox et al. 1990a,b,c) by Gaztañaga (1994) and Szapudi *et al.* (1995), (hereafter SDES), with generally good agreement with the Lick results of Szapudi, Szalay, & Boschán (1992), although there are some discrepancies. These may be due to differences in the scales over which these functions are averaged, but the differences between the catalogs or the measurement techniques cannot be precluded as the origin. Systematic variations in the measured magnitudes will induce artificial correlations, while different techniques will exhibit differing degrees of sensitivity to the sources of measurement error (Szapudi & Colombi 1996, hereafter SC96).

In this paper, we present an analysis of the higher order functions in the EDSGC catalog, an angular catalog covering 1000 square degrees (Heydon *et al.* 1989; Collins, Nichol & Lumsden 1992). We employ an efficient method based on factorial moments of cell counts. The exhaustive sampling of the catalog eliminates the measurement errors arising from the use of a finite number of sampling cells (SC96). This is verified on small angular scales using an algorithm equivalent to throwing at random an infinite number of cells to cover the catalog.

In the next section, we describe the EDSGC catalog, followed by an account of the measurement technique in §3. We present the results of the analysis in §4, and discuss their relation to previous analyses of other catalogs and to theoretical expectations in §5.

2. The Edinburgh/Durham Southern Galaxy Catalogue

The Edinburgh/Durham Southern Galaxy Catalogue (EDSGC) is a catalogue of 1.5 million galaxies cover $\simeq$ 1000 square degrees centered on the South Galactic Pole (SGP). The database was constructed from COSMOS scans (a microdensitometer) of 60 adjacent UK IIIa–J Schmidt photographic plates and reaches a limiting magnitude of $b_j = 20.5$. The entire catalogue has $< 10\%$ stellar contamination and is $\lesssim 95\%$ complete for galaxies brighter than $b_j = 19.5$ (Heydon *et al.* 1989). The two–point galaxy angular correlation function measured from the EDSGC has been presented by Collins, Nichol, & Lumsden (1992) and Nichol & Collins (1994).

A rectangular area of the catalog between $\alpha = 22^h$, passing through 0^h to 3^h, and declination $-42 \leq \delta \leq -23$, was suitable for our purposes. The original coordinates were converted to physical ones using an equal area projection. This did not affect the declination range, but to obtain a rectangular area the physical coordinates corresponding to right ascension, $\alpha - \alpha_{\rm min}\cos(\delta)$, were restricted to values less than 55°. This resulted in a sample of 2.9×10^5 galaxies, and a total effective survey area of 1045 square degrees, or $\simeq$ 997 square degrees after accounting for the cut-out regions.

Magnitude cuts were determined by practical considerations. The catalog is complete to about 20.3 magnitude. We adopt a limit half a magnitude brighter for our analysis to be conservative. To permit a direct comparison with results from the APM survey (Gaztañaga 1994), we used the magnitude cut $16.98 \leq m_{\rm EDS} \leq 19.8$. There is an offset in the magnitude scales of the two catalogs (Nichol 1992). Based on matching the surface densities listed in SDES, the magnitude range we have adopted corresponds approximately to the APM magnitude range $17 \leq m_{\rm APM} \leq 20$.

3. The method of analysis

The calculation of the higher order correlation functions consists of a sequence of three consecutive steps: estimation of the probability distribution, calculation of the factorial moments,

and extraction of the normalized, averaged amplitudes of the N-point correlation functions. We present the relevant definitions and theory below.

Let P_N denote the probability that a cell contains N galaxies, with implicit dependence on the cell size ℓ. The best estimator for P_N from the catalog is the probability that a randomly thrown cell *in the catalog* contains N galaxies (excluding edge effects, which are negligible for the scales in the present study, except perhaps on the largest scales as a result of the holes cut out around bright stars). This may either be calculated from the configuration of the points using a computer algorithm (see Szapudi 1996), or estimated by actually throwing cells at random,

$$\tilde{P}_N = \sum_{i=1}^{C} \delta(N_i = N), \tag{1}$$

where C is the number of cells thrown and N_i is the number of galaxies in cell i. It is desirable to use as many cells as possible, since for large C, the errors behave as (SC96)

$$E^{C,V} = (1 - \frac{1}{C})E^{\infty,V} + E^{C,\infty}, \tag{2}$$

where the $E^{C,V}$ is the total theoretical error (not including the sytematic errors of the catalog), $E^{\infty,V}$ is the 'cosmic' error associated with the finiteness of the catalog, and $E^{C,\infty}$ is the error associated with the finite number of cells used for the estimator. Since $E^{C,\infty} \propto C^{-1}$ (SC96), the lowest possible error is obtained for $C \longrightarrow \infty$. We employed such a code on scales smaller than 0.5°. On larger scales, up to 2°, we used massive oversampling (i.e. $C \gg V/v$, where V and v are the areas of the catalog and the cell respectively corresponding to a simple grid over the catalog), because of the limitations of the available computer resources for the current implementation of this memory intensive algorithm.

The factorial moments (see e.g. Szapudi & Szalay 1993), may be obtained from the probability distribution using

$$F_k = \sum P_N (N)_k, \tag{3}$$

where $(N)_k = N(N-1)..(N-1+k)$ is the k-th falling factorial of N. The F_k's directly estimate the moments of the underlying continuum random field which is Poisson sampled by the galaxies. This is equivalent to the ordinary moments after shot noise subtraction as can be seen from the relation with ordinary moments

$$\langle N^m \rangle = \sum_{k=0}^{m} S(m,k) F_k, \tag{4}$$

where $S(m,k)$ are the Stirling numbers of the second kind. The use of factorial moments simplifies all the expressions, since sums weighted by the Stirling numbers (shot noise) are eliminated. For instance, the factorial moments of a Poisson distribution are $F_k = \langle N \rangle^k$. These could have been obtained from a constant probability density $\delta(\epsilon - \langle N \rangle)$, which is the underlying continuum process. The ordinary moments of the Poisson distribution, however, will be more complicated, containing 'Poisson noise' from the previous equation.

The average of the N-point angular correlation functions on a scale ℓ is defined by

$$\bar{\omega}_N(\ell) = A(\ell)^{-N} \int d^2r_1 \ldots d^2r_N \omega_N(r_1, \ldots, r_N), \tag{5}$$

where ω_N is the N-point correlation function in the two dimensional survey, and $A(\ell)$ is the area of a square cell of size ℓ. We define s_N in the usual way,

$$s_N = \frac{\bar{\omega}_N}{\bar{\omega}_2^{N-1}}. \tag{6}$$

This definition is motivated by the assumed scale invariance of the N-point correlation functions in the strongly nonlinear limit (Balian & Schaeffer 1989),

$$\omega_N(\lambda r_1, \ldots, \lambda r_N) = \lambda^{-(\gamma-1)(N-1)} \omega_N(r_1, \ldots, r_N), \tag{7}$$

where γ is the slope of the three-dimensional two-point function. The coefficients also quantify the deviation from gaussian statistics, like skewness ($N = 3$) and kurtosis ($N = 4$). Derivations of the coefficients from perturbation theory have recently been performed in the weakly nonlinear limit for three dimensions by Juszkiewicz *et al.* (1993) and Bernardeau (1994a, b), and for two dimensions by Bernardeau (1995).

The factorial moments have an especially simple relation to the s_N's through the recursion relation (Szapudi & Szalay 1993),

$$s_k = \frac{F_k \bar{\omega}}{N_c^k} - \frac{1}{k} \sum_{q=1}^{k-1} \frac{(k-q) s_{k-q} F_q \binom{k}{q}}{N_c^q}, \tag{8}$$

where $N_c = \langle N \rangle \bar{\omega}$. Note that although the notation indicates a projected catalog, there are corresponding expressions for three dimensions.

The deprojection of the s_N's to their 3D counterparts has an intrinsic limitation due to the finite sizes of the cells. While the deprojection of any individual tree-structure is well-defined, care must be taken in interpreting the deprojected values of the cell-count determined s_N's, since these implicitly contain an averaging over trees within each cell (see SDES for a discussion). On small scales, where clustering is strongly nonlinear, the coefficients deproject to the 3D coefficients S_N defined by $S_N = \bar{\xi}_N / \bar{\xi}_2^{N-1}$, where the hierarchical assumption may be presumed valid. In this case,

$$s_N = R_N S_N, \tag{9}$$

where S_N is the corresponding three dimensional value for the spherically averaged $\bar{\xi}_N$'s. The projection coefficients R_N's are fairly insensitive to slight variations of the magnitude cut (see Table 2 in SDES), and the shape dependence is neglected according to the findings of Boschán, Szapudi, & Szalay (1994). We adopt the values of SDES. In the intermediate range of weakly nonlinear clustering, hierarchy-breaking terms become significant, and the differences between the conical averaging of the projected correlation functions and the spherical averaging of the three

dimensional functions become large (Bernardeau 1995). In this limit, the s_N deproject according to

$$s_N = R_N \Sigma_N, \tag{10}$$

where the Σ_N's involve averages only over the orthogonal parts of the wavevectors. (The expressions for R_N are identical in equations [9] and [10] for power-law power spectra.) Expressions for Σ_N for pure power-law power spectra have been worked out by Bernardeau (1995). For the depth of the EDSGC, the weakly nonlinear region corresponds to separations of $\theta \gtrsim 1°$.

4. Results

We measured counts in cells by throwing a large number of cells in the scale range of $0.015125° - 2°$ (corresponding to $0.1 - 13 h^{-1}$Mpc with $D \simeq 370 h^{-1}$Mpc, the approximate depth of the catalog). The largest scale is limited by the geometry induced by the cutout holes: the number of available cells would be severely limited for a measurement on significantly larger scales. The smallest scale approaches that of galaxy halos for the typical depth of the catalog. Note that even at the smallest scale, where the average count is only 0.0645 per cell, the s_N's are measurable to high accuracy because of the high oversampling and the efficient Poisson subtraction through the use of factorial moments. By comparison, the practice common in the literature is to stop at scales four times that at which the Poisson noise starts to dominate, i.e., where the average count approaches unity.

The results of both high and low oversampling measurements for P_N are displayed in Figure 1. The low sampling corresponds to covering the area with cells once only, i.e. the number of sampling cells is $C_V = V/v$, where V is the volume of the survey and v is the volume of the sampling cell at the given scale. As proved in (SC96), the 'number of statistically independent cells' , C^*, depends strongly both on scales and on the aims of the measurement, but for higher order statistics it is generally much larger than C_V. Therefore in the high oversampling measurements we doubled the sampling until the values of the s_N's, the primary goals of our measurement, stabilized. We also checked our results with 'infinite sampling' up to scales of 0.5°, where such a code could still work within our present computational limits. A comparison of the two curves shows the substantial improvement in accuracy achieved through oversampling.

Figure 2 shows the scale dependence of the s_N's determined from the counts in cells. The solid line corresponds to the measurements of the entire survey area with high oversampling. The dotted line is the same measurement with undersampling. For the error determination we divided the survey into four equal parts, similar to the approach of Gaztañaga (1994); it is likely that this procedure overestimates the cosmic error (SC96). The squares show the median of these measurements. The error bars are given by a robust determination of the dispersion: the spread between the minimum and maximum quartiles (see Hoaglin *et al.* 1983 for details). The error bars are shown only for those points for which there was sufficient physically valid data permitting a

determination of the quartiles. We note that on large scales the squares deviate from the solid line: this is presumably a result of edge effects. For s_3 and s_4, the errors range over $8-36\%$ and $19-56\%$, respectively. (We approximate 1σ errors by dividing the fourth spread errors by 1.349.) These may be compared with theoretical estimates for the errors. We base the estimates on the errors of the correlator moments over the entire catalog, according to SC96. For the first four moments, respectively, the errors are, ranging from large scales to small, $3-2\%$, $7-12\%$, $13-45\%$, and $23-63\%$. Although there is no simple formula relating the errors of the moments to the errors of the s_N's, it is likely that the errors at each order are dominated by the largest error; i.e., the highest moment. Thus, unless some cancelation effects are present, the last two values should well represent the errors on s_3 and s_4. These compare well with the empirical errors from the fourth spread above.

Figure 2 exhibits two plateaus, one at small scales ($< 0.03^\circ$) and a second at large ($> 0.5^\circ$). The large scale plateau is approaching the width of the survey, and so may merely reflect edge effects. The plateau at small separations, however, may indicate that the strongly nonlinear clustering limit has been reached, in which case the hierarchical form for the angular correlations should apply, for which the coefficients appear to converge. The values of s_N are provided in Table 1, and the ratios s_N/R_N in Table 2.

In order to probe more deeply into the weakly nonlinear regime, we performed a separate analysis extending to 4°. On these scales the majority of cells overlaps with some of the cut-out regions, therefore the analysis had to be done without the elimination of such cells, otherwise edge effects and cosmic errors from the resulting small area would have severly affected the measurement. After reanalysing all scales without eliminating cells containing the cut out holes, we concluded the effect is a slight low bias, if any, which is well within the statistical errors. We nonetheless find good agreement with the smaller scale analysis for $\theta \leq 1^\circ$. We obtain in the larger scale analysis $s_3 = 5.55$ and $s_4 = 58.0$ at $\theta = 2^\circ$, and $s_3 = 6.88$ and $s_4 = 48.2$ at $\theta = 4^\circ$. These correspond to $\Sigma_3 = 4.71$ and $\Sigma_4 = 39.1$ at $13h^{-1}$ Mpc separation, and $\Sigma_3 = 5.84$ and $\Sigma_4 = 32.5$ at $26h^{-1}$ Mpc separation.

5. Discussion

We find generally good agreement between the amplitudes s_N in the EDSGC and those derived by Gaztañaga for the APM survey. While the values we derive tend to lie systematically above those for the APM (Figure 3), each agrees within the error estimate. Unlike Gaztañaga's results, however, we find that the s_N's generally reach a plateau at small scales when we sample the catalog exhaustively (Figure 2). By contrast, Gaztañaga finds that the coefficients tend to turn over at small separations. Without exhaustive sampling, we obtain a similar result, demonstrating the need to sample the catalog many times to obtain an accurate determination of the clustering strength on small scales. The reason for the difference from Gaztañaga's results is unclear, but the difference may account for the higher values for the coefficients we obtain from the EDSGC.

A plateau at small separations may be expected when the clustering becomes strongly nonlinear. The effect is found, for example, in the N-body experiments for scale-free clustering by Colombi, Bouchet, & Hernquist (1995). If the clustering we measure is strongly nonlinear on the smallest scales, then we are permitted to identify $S_N = s_N/R_N$ in Table 2 at small separations. We may then in turn derive the 3D clustering amplitudes Q_N. These are defined within the hierarchical model

$$\xi_N(r_1,\ldots,r_N) = \sum_{k=1}^{K(N)} Q_{Nk} \sum^{B_{Nk}} \prod^{N-1} \xi(r_{ij}), \tag{11}$$

where $\xi(r) \equiv \xi_2(r) = (r/r_0)^\gamma$, as the average of the Q_{Nk}

$$Q_N = \frac{\sum_{k=1}^{K(N)} Q_{Nk} B_{Nk} F_{Nk}}{N^{(N-2)}}, \tag{12}$$

where F_{Nk} are the form factors associated with the shape of cell of size unity (see Boschán, Szapudi, & Szalay 1994 for details)

$$F_{Nk} = \int_1 d^3q_1 \ldots d^3q_N \prod^{N-1} \left\{ \mid q_i - q_j \mid^\gamma \int_1 d^3p_1 d^3p_2 |p_1 - p_2|^{-\gamma} \right\}^{-1}. \tag{13}$$

The product above runs over the $N-1$ edges of a tree. The summation in equation (11) is over all possible N^{N-2} trees with N vertices. In the sum, every $\xi(r_{ij})$ corresponds to an edge $r_{ij} = \mid r_i - r_j \mid$ in a tree spanned by $r_1,\ldots,r_N$. For each tree, there is a product of $N-1$ two-point functions and a summation over all the B_{Nk} labelings of all the $K(N)$ distinct trees.

Using the values for $r = 0.1$ Mpc in Table 2, we find for $N = 3-9$, $Q_N = (2.0, 7.2, 28, 98, 270, 570, 850)$. The values for Q_3 and Q_4 somewhat exceed those found for the Lick-Zwicky sample (Groth & Peebles 1977; Fry & Peebles 1978; Szapudi, Szalay & Boschán 1992), and greatly exceed the values found for the CfA1 and SSRS surveys (Gaztañaga 1992). The discrepancy between the larger angular samples and the smaller samples used for the redshift surveys has been previously noted by Fry & Gaztañaga (1994). Our results suggest the discrepancy at small scales may be even larger. The reason for the difference in the values is unknown, but may be a result of cosmic variance. It appears not to be a result of the added redshift information, since Gaztañaga (1994) found that suppressing the redshifts in the CfA1 and SSRS surveys and treating the samples as angular catalogs had little effect on the values.

In the limit of weakly nonlinear clustering, it is possible to compare the clustering coefficients with theoretical predictions for a given power spectrum (Juszkiewicz, Bouchet, & Colombi 1993; Bernardeau 1994a,b; Bernardeau 1995). We find from the values of Σ_3 and Σ_4 at separations of 6.5, 13, and 26 h^{-1} Mpc the values $n_{\rm eff} = (-1.2, -1.7, -2.5)$ for $N = 3$, and $n_{\rm eff} = (-1.3, -1.6, -1.3)$ for $N = 4$, using the results of the larger 4° analysis for $\theta > 1^\circ$ from the previous section, and the expressions relating Σ_N to n in Bernardeau (1995). The trend of decreasing $n_{\rm eff}$ with increasing scale is suspect. For a power spectrum like CDM, $n_{\rm eff}$ increases with increasing scale. The inverse

trend may indicate that edge effects are significant on these scales and are compromising the determination of s_N on scales exceeding 1°.

While the weakly nonlinear limit should break down on scales smaller than 1°, it is informative to explore the inferred dependence of $n_{\rm eff}$ on scale to smaller values. Colombi et al. (1995) find from N-body experiments for scale-free clustering that the values for S_N vary only slightly with scale, increasing for small separations where the clustering becomes strongly nonlinear. They find, independent of n,

$$S_N \simeq [D(\bar{\xi}_2)]^{N-2}\, \tilde{S}_N, \tag{14}$$

for $N = 3$, 4, and 5, where $D(\bar{\xi}_2) = (\bar{\xi}_2/100)^{0.045}$ and $\tilde{S}_N$ is the value of S_N for $\bar{\xi}_2 = 100$. The dependence on scale implies a weak departure from the hierarchical clustering behavior, since the S_N depend on scale. The dependence is so weak, however, that the departure is slight. We compare the clustering amplitudes found in the EDSGC with this relation in Figure 4a. The agreement in the strong clustering limit is remarkably good, particularly for $N = 3$ and 4. Because we have only angular information, it is not possible to determine whether the deviation from the scaling relation for $\log \bar{\xi}_2 < 2$ is a real effect or a consequence of the inherent limitations of extracting 3D information from a projected catalog. It should be noted that the agreement is particularly remarkable since hierarchical clustering is assumed for the underlying distribution in order to convert from the projected correlations to the 3D, while the relation of equation (14) violates this assumption. This suggests that the hierarchical model is a good, though perhaps not perfect, description of the clustering.

Colombi et al. find that the clustering for $N = 3$, 4, and 5 may be described by a single effective spectral index $n_{\rm eff}$, found from the expressions for weakly nonlinear clustering (Juszkiewicz, Bouchet, & Colombi 1993; Bernardeau 1994a,b). Although the relations between the S_N and n from weakly nonlinear theory do not apply for strong clustering, and even less so in an angular catalog, we may adopt them to obtain a formal value for $n_{\rm eff}$ as done by Colombi et al. (1995, 1996). We do so by fitting s_N/R_N to the expressions for S_N in the limit of weakly nonlinear clustering for $N = 3 \ldots 6$ using a least squares method, for $\theta \leq 1^\circ$. The results are shown in Figure 4b, including the values derived for each N individually. Within the error estimates, a single value of $n_{\rm eff}$ appears to provide an adequate description of the clustering amplitudes, although the errors are large for small separations. A comparison with N-body results for scale-dependent clustering models, like a CDM-dominated cosmology, could be very illuminating.

I.S. thanks S. Colombi, J. Frieman, and A. Szalay for stimulating discussions. S. Colombi provided the theoretical error estimates in §4. I.S. was supported by DOE and NASA through grant NAG-5-2788 at Fermilab. A.M. is grateful to the W. Gaertner Fund at the University of Chicago for support.

6. Figure Captions

Figure 1. Shows the distributions P_N of counts in cells measured in the EDSGC catalog. The solid line corresponds to exhaustive sampling, while the dotted line to severe undersampling. The curves from left to right correspond to cell sizes from 0.015125° doubling up to 2°. Exhaustive sampling is essential on scales exceeding $\sim 0.2^\circ$.

Figure 2. The solid line is the measurement of the s_N's over the entire survey area with exhaustive sampling, the dotted line is the same with undersampling. Undersampling results in a systematic underestimate of the coefficients. The squares show the median of the measurements in four equal parts of the survey, and the errors are calculated from the fourth spread. The misalignment of the squares and the solid line at the largest scales may be a result of edge effects. The triangles show the Σ_N's corresponding to the best fitting formal $n_{\rm eff}$ (see text).

Figure 3. A comparison with the results from the APM survey. The points with short error bars are the values determined for the APM catalog by Gaztañaga 1994, which were averaged over the range 0.09° to 0.53°. The squares (with the larger error bars), and the triangles are our measurements with exhaustive sampling and low sampling, respectively, averaged over the same range of scales. The three points were shifted slightly for display purposes. Note that the range quoted is adjusted for the equivalent size of a square cell, thus differing from the scales in Gaztañaga (1994) by a factor of $\pi^{1/2}$.

Figure 4. a. The clustering amplitudes s_N/R_N as a function of the average 2-point function $\bar{\xi}_2$. Also shown is the scaling relation of Colombi et al. (1995) found in the strongly nonlinear limit in N-body experiments with scale-free initial conditions. b. The best formal fits for $n_{\rm eff}$ (*solid*), using up to sixth order quantities. Also shown are the values determined from each N separately, including an indication of the errors based on the upper and lower quartile values for each S_N. Within the errors, the clustering may be described by a single value of $n_{\rm eff}$. Shown are the values of $n_{\rm eff}$ for $N = 3$ (dotted), $N = 4$ (*short-dashed*), $N = 5$ (*long-dashed*), and $N = 6$ (*dot-dashed*).

REFERENCES

Balian, R., & Schaeffer, R. 1989, A&A, 220, 1

Bernardeau, F. 1994a, ApJ, 433, 1

Bernardeau, F. 1994b, A&A, 291, 697

Bernardeau, F. 1995, A&A, 301, 309

Boschán, P., Szapudi, I., & Szalay, A. 1994, ApJS, 93, 65

Bouchet, F.R., Strauss, M.A., Davis, M., Fisher, K.B., Yahil, A., & Huchra, J.P. 1993, ApJ, 417, 36

Collins, C. A. Nichol, R. C., & Lumsden, S. L. 1992, MNRAS, 254, 295

Colombi, S., Bernardeau, F., Bouchet, F.R., & Hernquist, L. 1996, (in preparation)

Colombi, S., Bouchet, F.R., & Hernquist, L. 1995, A&A, 281, 301

Colombi, S., Bouchet, F.R., & Schaeffer, R. 1994, A&A, 281, 301

Fry, J. N. & Gaztañaga, E. 1994, ApJ, 425, 1

Fry, J., & Peebles, P.J.E. 1978, ApJ, 221, 19

Gaztañaga, E. 1992, ApJ, 319, L17

Gaztañaga, E. 1994, MNRAS, 268, 913

Groth, E.J., & Peebles, P.J.E. 1977, ApJ, 217, 385

Hamilton, A. J. S. 1988, ApJ, 332, L67

Heydon-Dumbleton, H. H., Collins, C. A., & MacGillivray, H. T. 1989, MNRAS, 238, 379

Hoaglin, D.C., Mosteller, F., & Tukey, J.W. (eds.) 1983, Understanding Robust and Exploratory Data Analysis (New York: John Wiley & Sons)

Juszkiewicz, R., Bouchet, F. R., & Colombi, S. 1993, ApJ, 412, L9

Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, L. 1990a, MNRAS, 242, 43P

Maddox, S. J., Sutherland, W. J., Efstathiou, G., & Loveday, L. 1990b, MNRAS, 243, 692

Maddox, S. J., Sutherland, W. J., Efstathiou, G., & Loveday, L. 1990b, MNRAS, 246, 433

Meiksin, A., Szapudi, I., & Szalay, A., 1992, ApJ, 394, 87

Nichol, R. C. 1992, PhD thesis, University of Edinburgh

Nichol, R. C., & Collins, C. A. 1994, MNRAS, 265, 867

Peebles, P.J.E. 1980, The Large Scale Structure of the Universe (Princeton: Princeton University Press)

Peebles, P.J.E. 1993, Principles of Physical Cosmology (Princeton: Princeton University Press)

Shane, C.D., & Wirtanen, C.A. 1967 Publ. Lick. Obs., 22, part 1

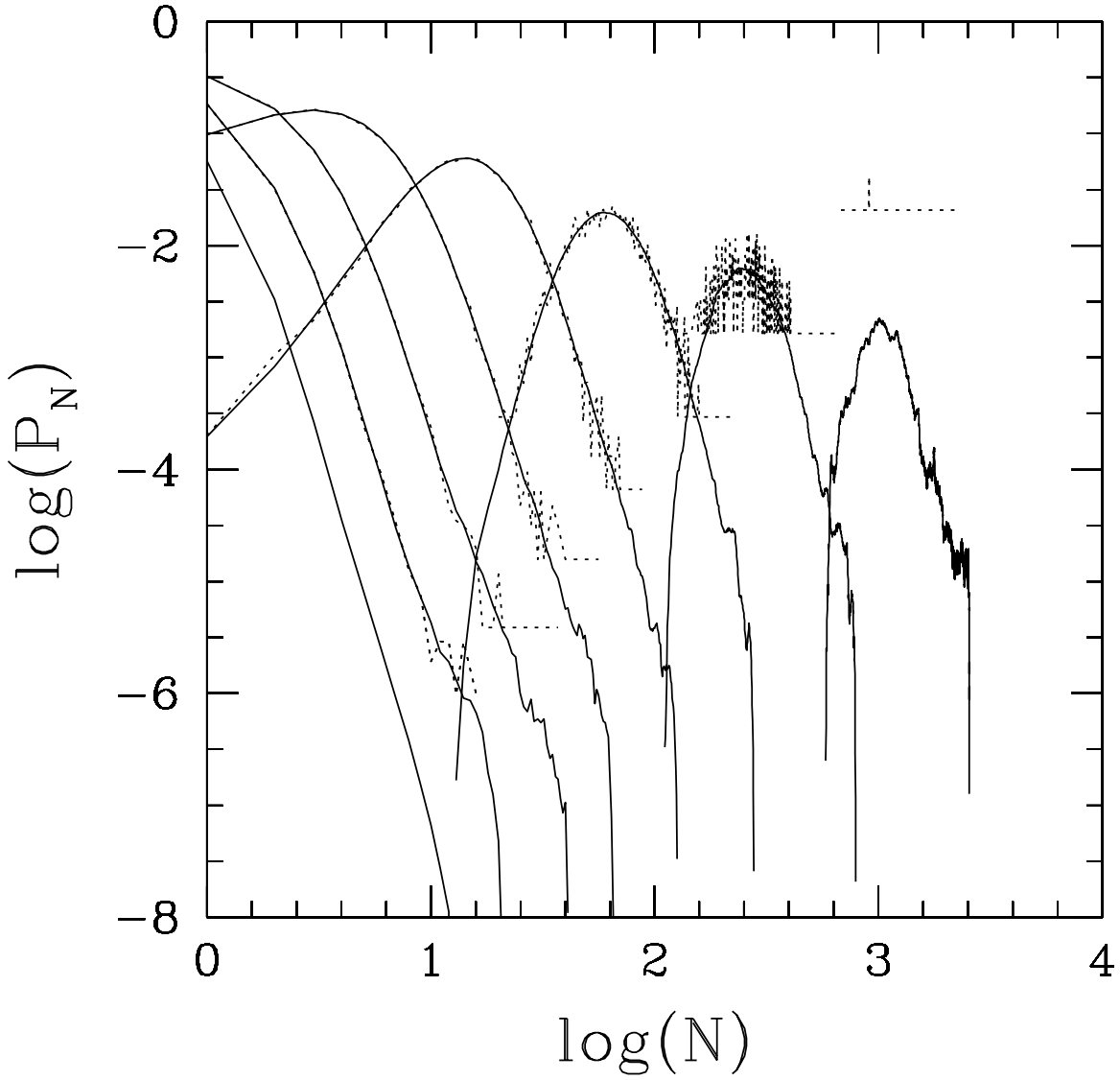
Szapudi, I. 1996, (in preparation)

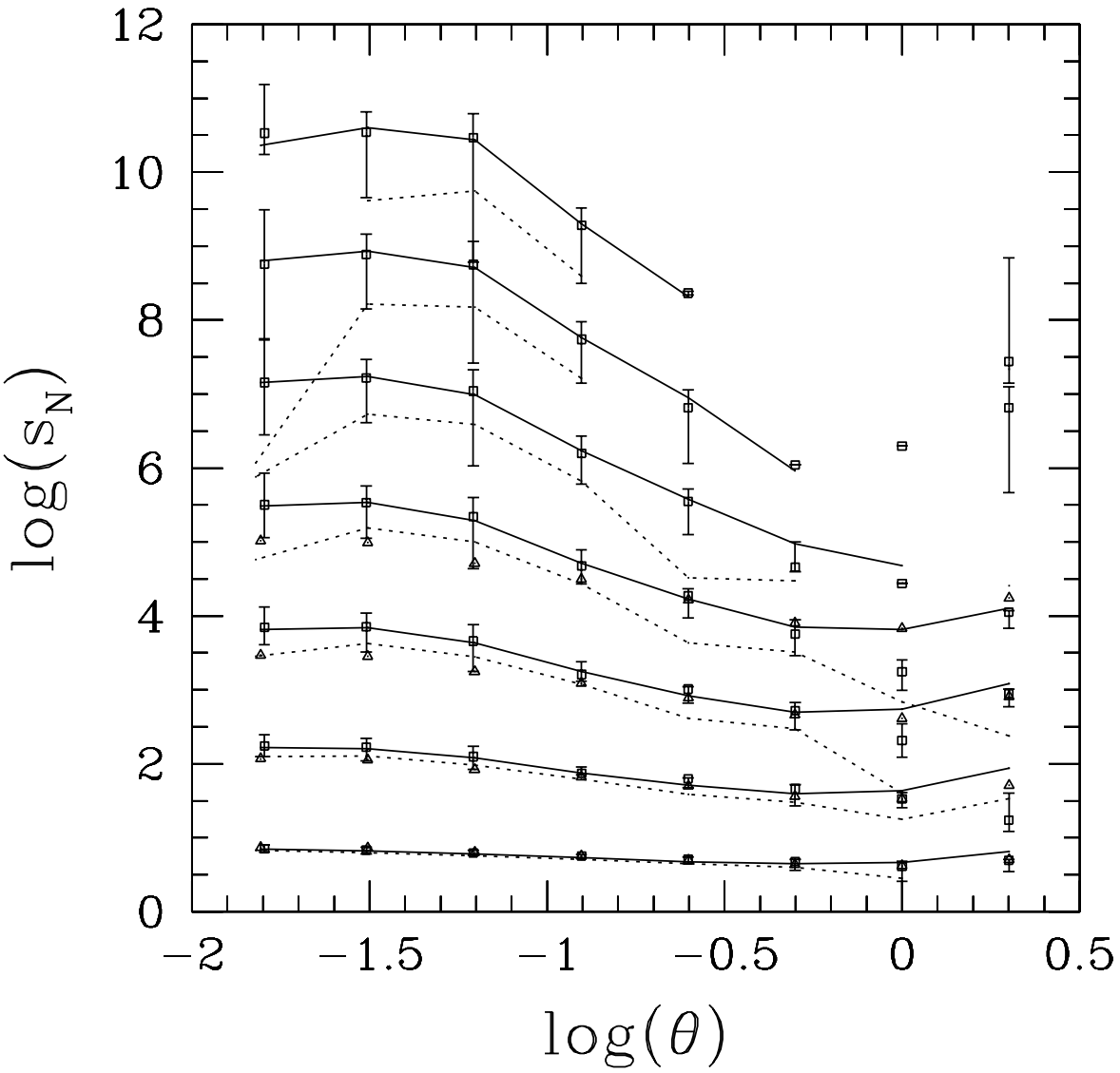
Szapudi, I., & Colombi, S. 1996, ApJ (submitted) (SC96)

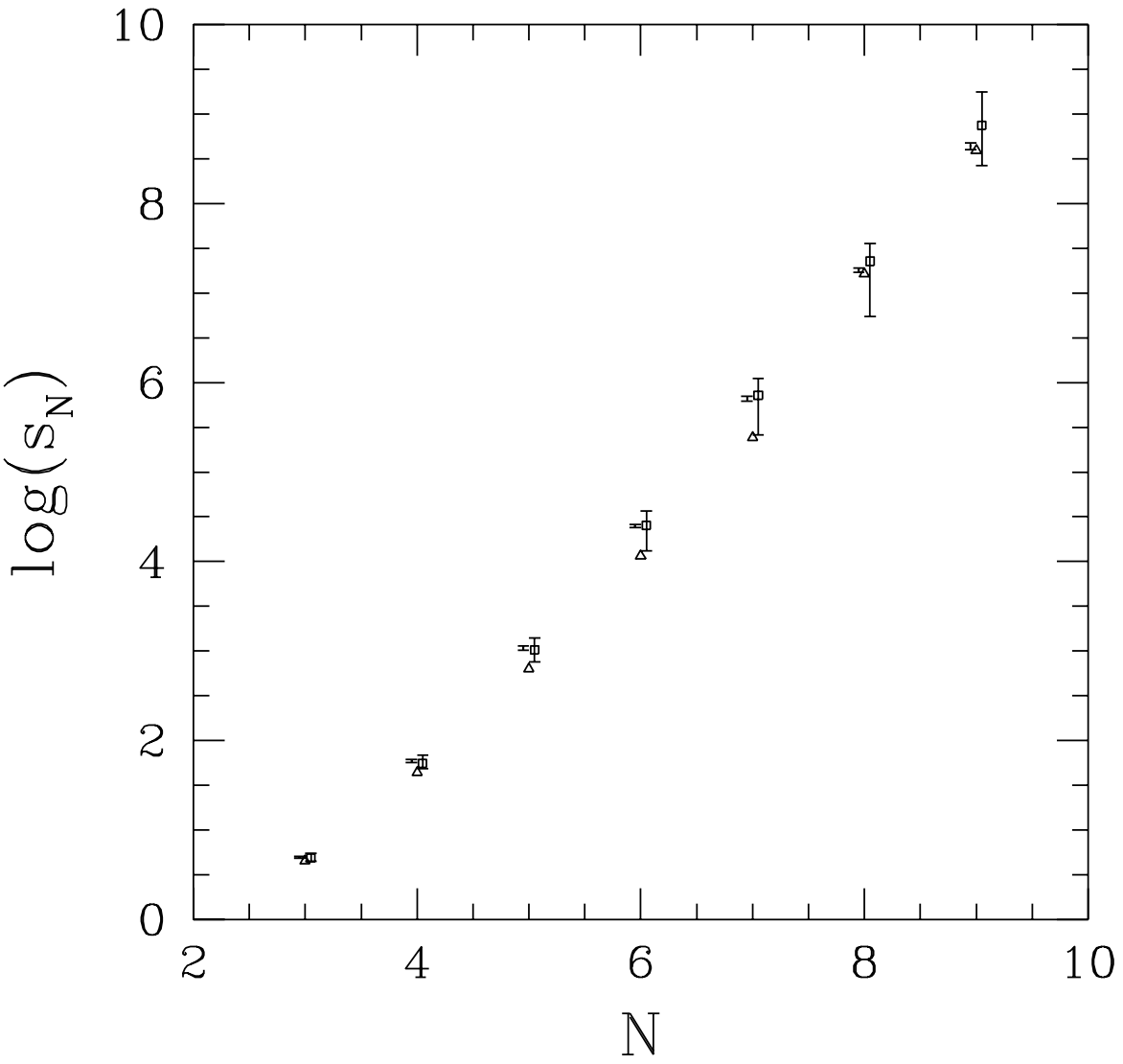
Szapudi, I., Dalton, G., Efstathiou, G.P., & Szalay, A. 1995, ApJ, 444, 520 (SDES)

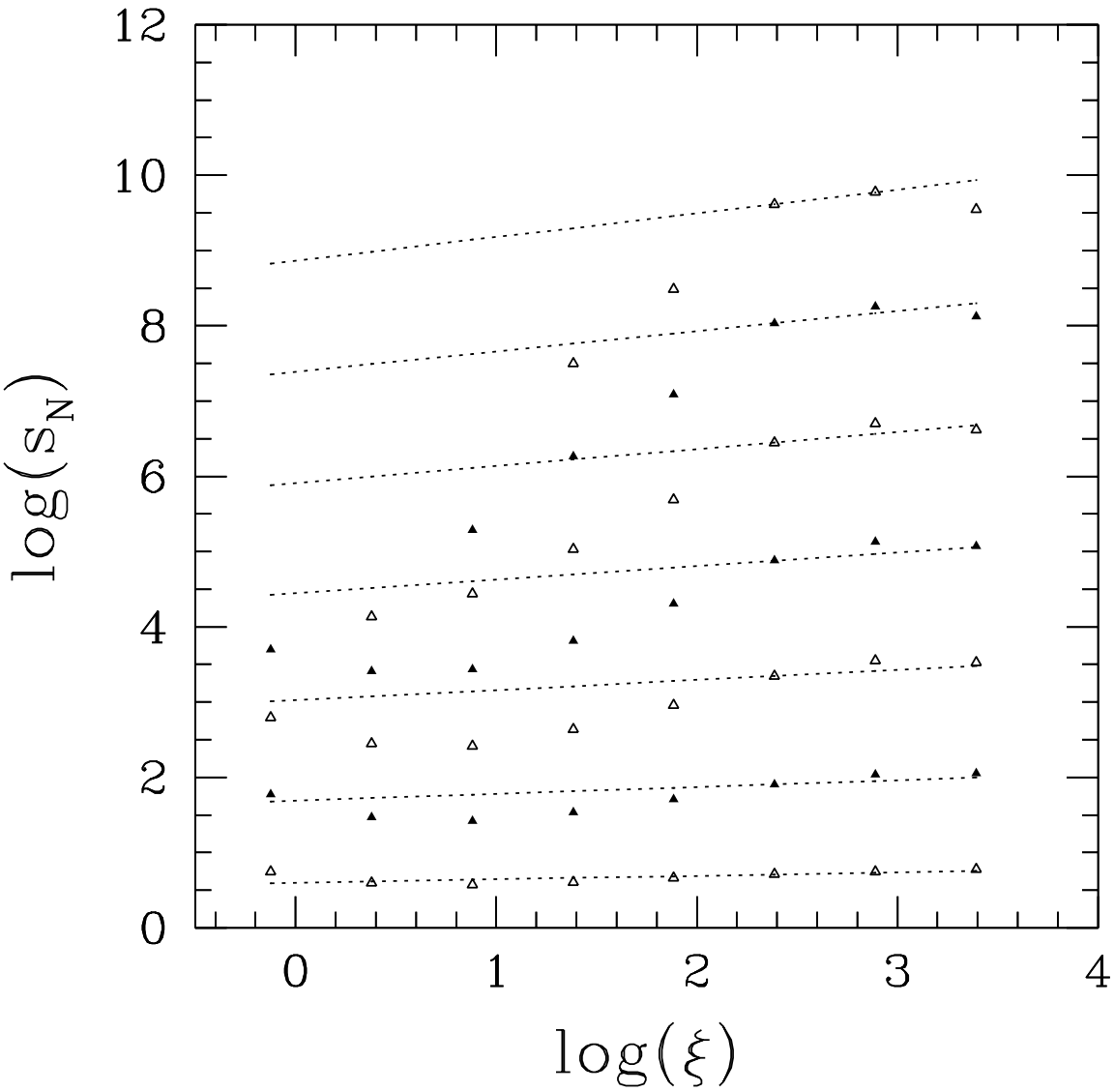
Szapudi, I. & Szalay, A. 1993, ApJ, 408, 43

Szapudi, I., Szalay, A., & Boschán, P. 1992, ApJ, 390, 350


0
−2
−4
−6
−8
0
1
2
3
4
log(P_N)
log(N)


12
10
8
6
4
2
0
−2
−1.5
−1
−0.5
0
0.5
log(S_N)
log(θ)


10
8
6
4
2
0
2
4
6
8
10
N
log(s_N)


12
10
8
6
4
2
0
0
1
2
3
4
log(s_N)
log(ξ)

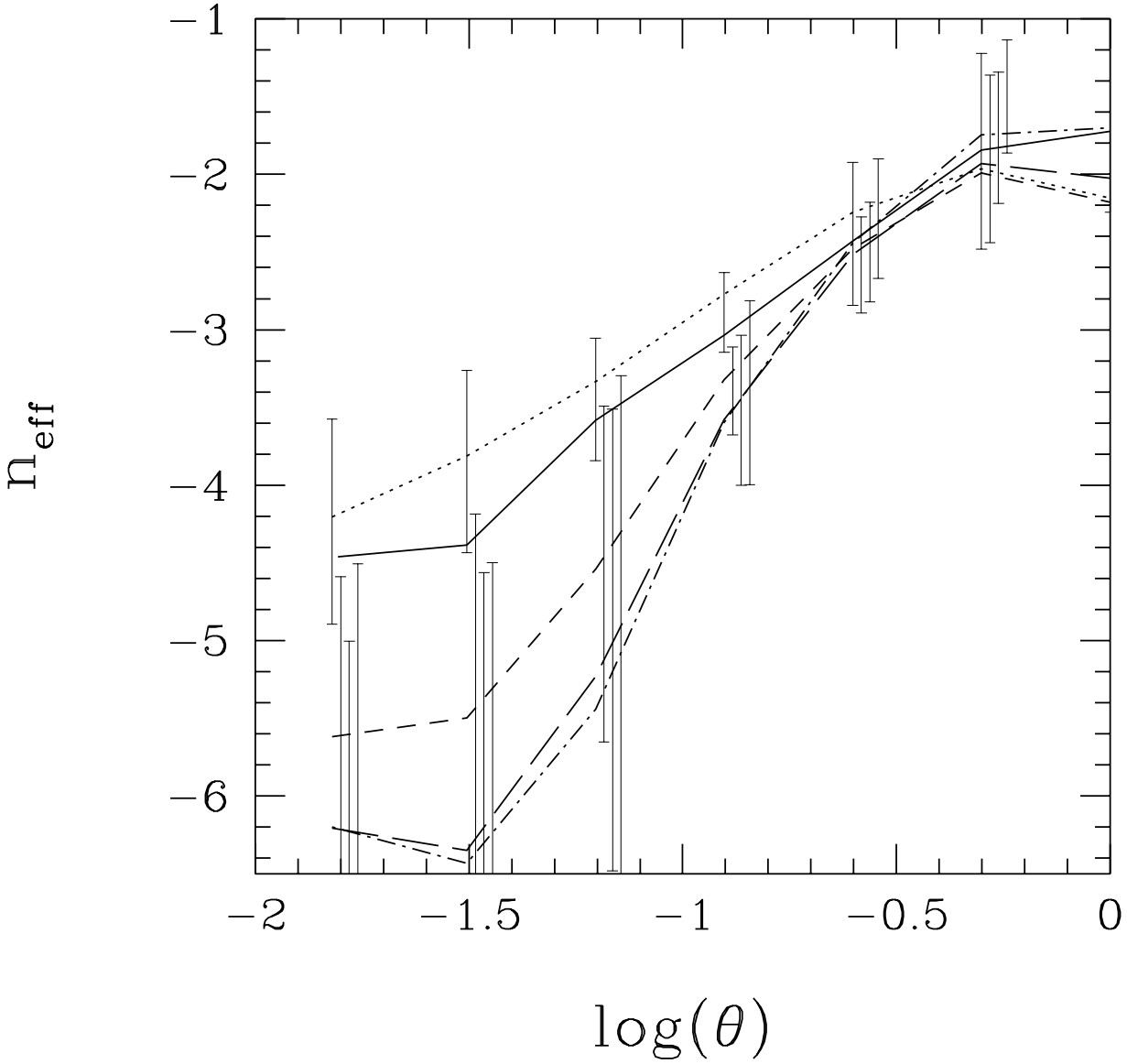

−1
−2
−3
−4
−5
−6
−2
−1.5
−1
−0.5
0
$n_{\rm eff}$
$\log(\theta)$

TABLE 1

VALUES OF s_N

θ (°)	s_3	s_4	s_5	s_6	s_7	s_8	s_9
2	6.55	87.4	1.21E3	1.28E4	...	...	...
1	4.65	43.7	547	6.60E3	4.78E4	...	...
0.5	4.43	39.3	501	7.04E3	9.50E4	9.19E5	...
0.25	4.76	51.1	832	1.67E4	3.76E5	8.84E6	2.06E8
0.125	5.37	75.7	1.76E3	5.19E4	1.70E6	5.83E7	2.02E9
0.0625	6.02	119	4.29E3	1.95E5	9.83E6	5.16E8	2.73E10
0.03125	6.58	161	6.93E3	3.45E5	1.75E7	8.58E8	3.98E10
0.015125	7.03	166	6.54E3	3.04E5	1.44E7	6.30E8	2.32E10

TABLE 2

VALUES OF s_N/R_N

r (Mpc)	s_3/R_3	s_4/R_4	s_5/R_5	s_6/R_6	s_7/R_7	s_8/R_8	s_9/R_9
13.0	5.64	61.0	662	5.35E3	...	...	...
6.5	4.01	30.5	299	2.76E3	1.51E4	...	...
3.2	3.82	27.4	274	2.95E3	3.01E4	2.18E5	...
1.6	4.10	35.7	455	7.02E3	1.19E5	2.09E6	3.63E7
0.81	4.63	52.8	963	2.18E4	5.38E5	1.38E7	3.55E8
0.40	5.19	83.2	2.34E3	8.19E4	3.12E6	1.22E8	4.79E9
0.20	5.67	112	3.78E3	1.44E5	5.54E6	2.03E8	6.99E9
0.10	6.06	116	3.57E3	1.27E5	4.57E6	1.49E8	4.08E9